Assignment Report
ET 4389 : Complex Networks - from Nature to Man-made Networks

Group 1

Alon Dolev (4251873)
Mani Prashanth Varma Manthena (4243846)

1 Topological Properties of Dutch Soccer - Real World Network

1.1) Dutch Soccer - Real World Network

Number of nodes (*N*) = **685**
Number of links (*L*) = **10310**
Link density (*p*) = **0.044**
Average degree (*E*[D]) = **30.102**
Degree variance (*Var*[D]) = **449.916**

Degree distribution (Linear-linear scale plot) :

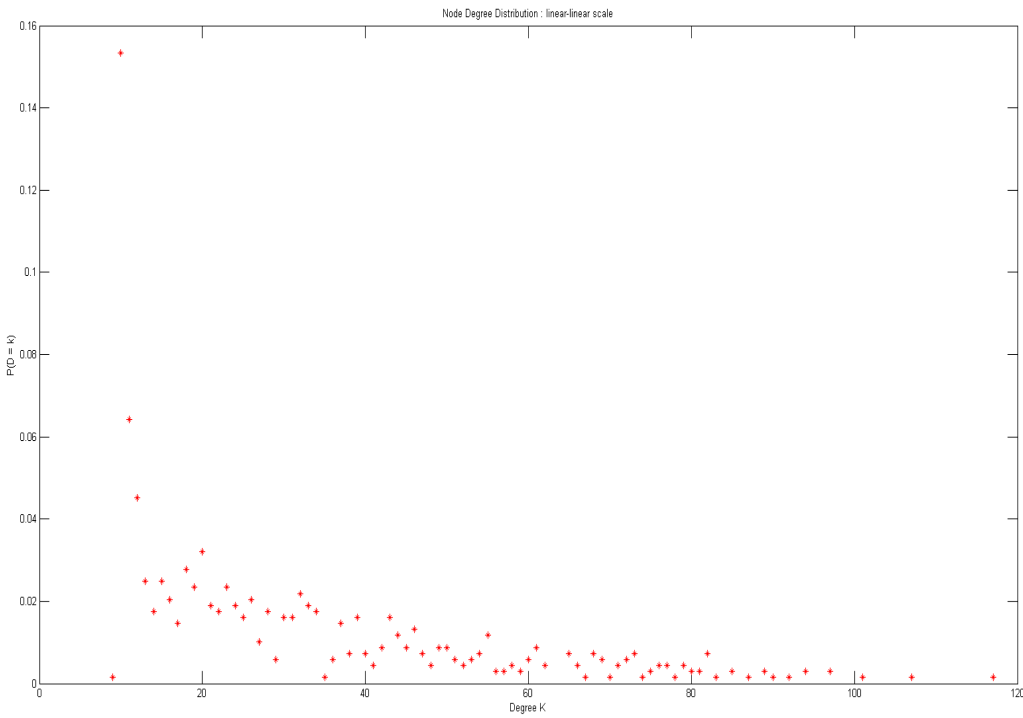

Fig. 1: Degree distribution of Dutch Soccer - Real World Network in Linear-linear scale

Yes, It follows power law distribution (Pr[D = K] ~ cK^{γ}) with power exponent (γ) = **-1.256** and power law constant (c) = **0.796**

Degree distribution (Log-log scale plot) with fitting curve :

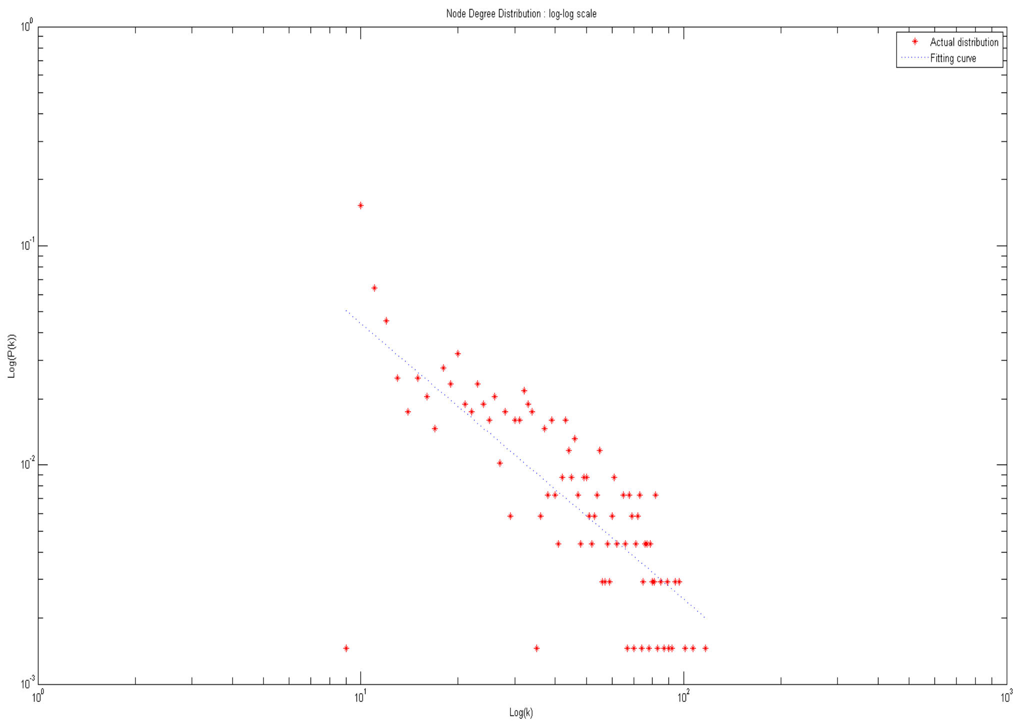

Fig. 2**:** Degree distribution of Dutch Soccer - Real World Network in Log-log scale with fitting curve

1.2) Degree correlation (assortativity) (ρ_D) = **- 0.0633**

Physical Meaning: As rou D < 0, the network is (degree) disassortative which means that high degree nodes in the network tend to link/connect with other low degree nodes in the network or vice versa i.e. presence of links/connections between nodes of different degrees.

1.3) Clustering coefficient (*C*)= **0.750**

1.4) Average hop count of the shortest paths between all node pairs (*E*[H]) = **4.458**

Diameter of the Dutch Soccer - Real World Network (*H* max) = **11**

1.5) Largest eigenvalue (spectral radius) of the adjacency matrix (λ_1) = **50.842**

1.6) Second smallest eigenvalue (algebraic connectivity) of the Laplacian matrix (μ_{N-1}) = **0.161**

2 Effect of Degree Correlation on Network Properties

2.1) Average of each network metric over 100 network instances formed by applying degree preserving rewiring to the Dutch Soccer - Real World Network with degree correlation $\rho \; \varepsilon$ [0.3, 0.31]

(ρ_D) = **0.3001**

C = **0.193**

E[H] = **2.400**

H max = **4**

λ_1 = **51.385**

μ_{N-1} = **5.593**

2.2) Average of each network metric over 100 network instances formed by applying degree preserving rewiring to the Dutch Soccer - Real World Network with degree correlation $\rho \; \varepsilon$ [-0.31, -0.3]

(ρ_D) = **-0.3001**

C = **0.214**

E[H] = **2.245**

H max = **4**

λ_1 = **42.887**

μ_{N-1} = **7.240**

2.3) Comparison of the network metrics in the Dutch Soccer - Real World Network and in the two classes of networks formed by applying degree preserving rewiring to the Dutch Soccer - Real World Network as in sections 2.1 and 2.2

Network Metric	Actual Network (Dutch Soccer)	Assortative Rewired Network with $\rho \ \varepsilon [0.3, 0.31]$	Disassortative Rewired Network with $\rho \ \varepsilon [-0.31, -0.3]$	Effect of Rewiring (Degree Correlation)
Degree correlation coefficient (ρ_D)	**- 0.0633**	**0.3001**	**-0.3001**	assortative rewiring $\rho_D > 0$ and disassortative rewiring $\rho_D < 0$
Clustering coefficient (*C*)	**0.750**	**0.193**	**0.214**	rewiring reduced *C* value
Average hop count (*E*[H])	**4.458**	**2.400**	**2.245**	rewiring reduced *E*[H] value
Diameter (*H* max)	**11**	**4**	**4**	rewiring reduced *H* max value
Spectral radius (λ_1)	**50.842**	**51.385**	**42.887**	assortative rewiring increased λ_1 value whereas disassortative rewiring reduced it
Algebraic connectivity (μ_{N-1})	**0.161**	**5.593**	**7.240**	rewiring increased μ_{N-1} value

Degree correlation coefficient (ρ_D) : range : $-1 \leq \rho_D \leq 1$

Physical meaning: Higher values of ρ_D (i.e. $\rho_D > 0$) means that the network is degree assortative where nodes tend to attach/link to other nodes with similar degree e.g. online social networks and whereas in the case of lower values of ρ_D (i.e. $\rho_D < 0$) means that the network is degree disassortative where high degree nodes in the network tend to link/connect with other low degree nodes in the network or vice versa i.e. presence of links/connections between nodes of different degrees e.g. biological networks. Networks with $\rho_D < 0$ are more vulnerable to both random and targeted attacks.

Clustering coefficient (*C*) : range : $0 \leq C \leq 1$

Physical meaning: It determines the scale to which nodes in the network tend to form clusters. Higher values of *C* result in formation of hubs (i.e. high degree nodes) as in scale free networks which are robust against random failures but not so robust in the case of targeted attacks.

Average hop count (*E*[H]) :

Physical meaning: It is the average of all the shortest paths between all the possible node pairs in the network. Further, smaller values of average hop count means that all the nodes in the network are closely connected/separated to each other. Thus, such networks have the advantage in terms of bandwidth/cost savings (i.e. efficiency of transport on the network) and at the same time they are more robust to any kind of failures/attacks as the network is still operational after failure of few nodes in the network.

Diameter (*H* max) :

Physical meaning: It is the maximum of all the shortest paths between all the possible node pairs in the network. Lower values of *H* max may corresponds to networks which are robust against random failures but not so robust in the case of targeted attacks e.g. scale free networks.

Spectral radius (λ_1) :

Physical meaning: It determines robustness of a network against both failures and spread of epidemics. Epidemic threshold is inversely proportional to λ_1. So, networks with lower values of λ_1 are more robust against spread of epidemics. Further, larger values of λ_1 correspond to smaller values of diameter which in turn corresponds to higher robustness in the network against both node and link failures.

Algebraic connectivity (μ_{N-1}) :

Physical meaning: It determines the connectivity and number of disjoint clusters in the network. It also measures how difficult it is to break the network into disjoint clusters or components. Higher the value of μ_{N-1}, greater is the robustness of the corresponding network to both node and link failure/removal.

Design recommendation for the communication network :

Taking all the network metrics and their corresponding values into account from the above table. It can be said that there is no single best design from the three possible designs for the communication network.Choice of the design for the communication network depends on its application and the trade offs that can be made in the communication network.

Importance must be given to spectral metrics i.e. λ_1 and μ_{N-1} over other network metrics as they are essential in network characterizations (*Refer metric correlation in the lecture slides).

Firstly, All communication networks in general require and demand high efficiency of data/information transport on the network. To achieve this the corresponding network design must have lower value of average hop count (*E*[H]). Looking at the above table, the rewired networks have lower average hop count (*E*[H]) compared to the actual network i.e. Dutch Soccer. Average hop count (*E*[H]) is almost similar in both the rewired networks but disassortative rewired network has slightly lower value which makes it a bit advantageous design choice. So, on a broad scale the contenders for the communication network design are the two rewired networks i.e. assortative rewired network with $\rho\ \varepsilon$ [0.3, 0.31] and disassortative rewired network with $\rho\ \varepsilon$ [-0.31, -0.3].

Secondly, the choice between the two rewired networks depends upon the application for which the communication network will be used. If the communication network is used for commercial applications like internet, mobile, TV services which require and demand continuous coverage/service even in the case of random failure of (fewer) intermediate nodes/routers in the network which generally happens in such networks. In such a case, the disassortative rewired network is a better design choice for the communication network as it has higher value of μ_{N-1} which translates to greater robustness to such random failures than its counterpart and at the same time it has lower value of λ_1 which translates to higher robustness against spread of epidemics/viruses in the network compared to its counterpart. On the other hand, for non-commercial applications like defence, military and other highly confidential services which require high security on the data/information in the network and as such those networks are always prone to constant targeted attacks. In such networks, the assortative rewired network is a better design choice for the communication network as it has higher value of ρ_D which translates to higher robustness towards both random and targeted attacks compared to its counterpart.

Recommendation : Disassortative rewired network with $\rho\ \varepsilon$ [-0.31, -0.3] is a better design choice for the communication network **(Trade off : Robustness of the network against targeted attacks)**

3 Network Models

3.1) Erdos-Renyi (ER) random graphs :

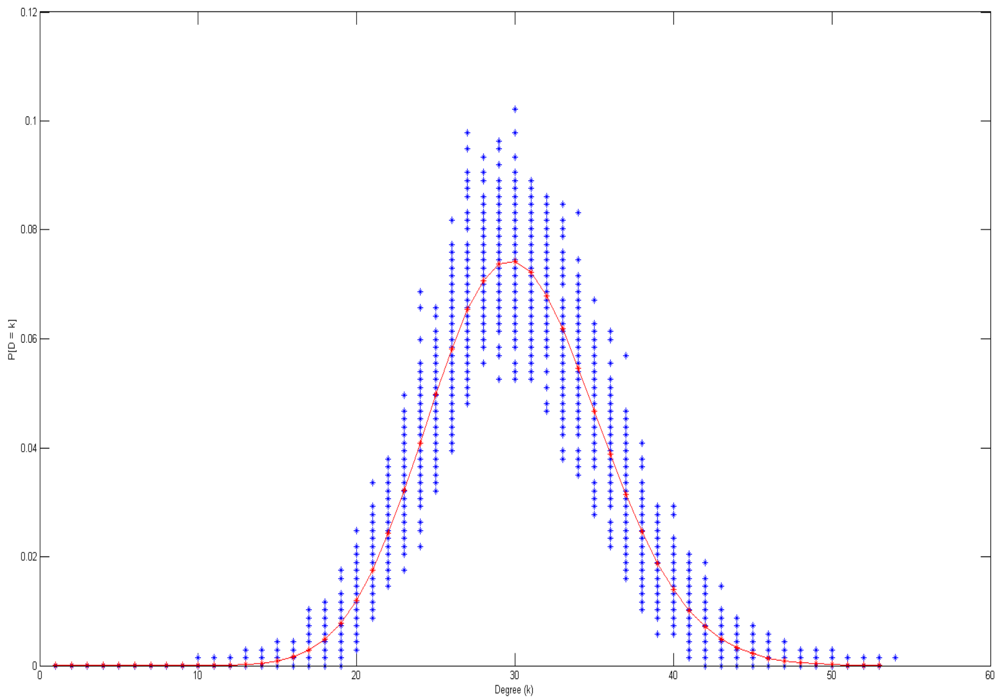

Fig. 3: Degree distribution of 100 Erdos-Renyi random graphs with number of nodes (N) = 685 and link density (p) = 0.044. Expected theoretical distribution in red

3.2) Barabasi-Albert (BA) scale free networks :

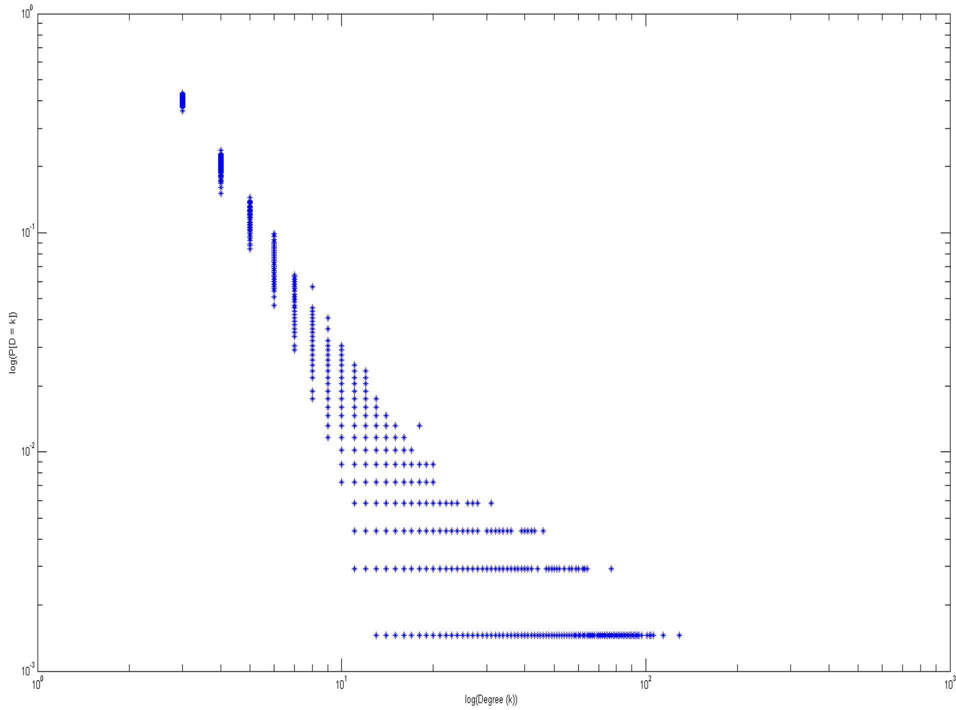

Fig. 4: Degree distribution of 100 scale-free random graphs generated using the Barabasi-Albert algorithm with number of nodes (N) = 685

The power exponent of γ = **-1.822** was determined by calculating the mean of all the highest exponents of 2-D *polyfit* function (curve fitting) applied over all the 100 generated networks.

4 Percolation and Network Robustness

*Note: All the plots below are average over 100 iterations/instances with random node failures/removal.

4.1)

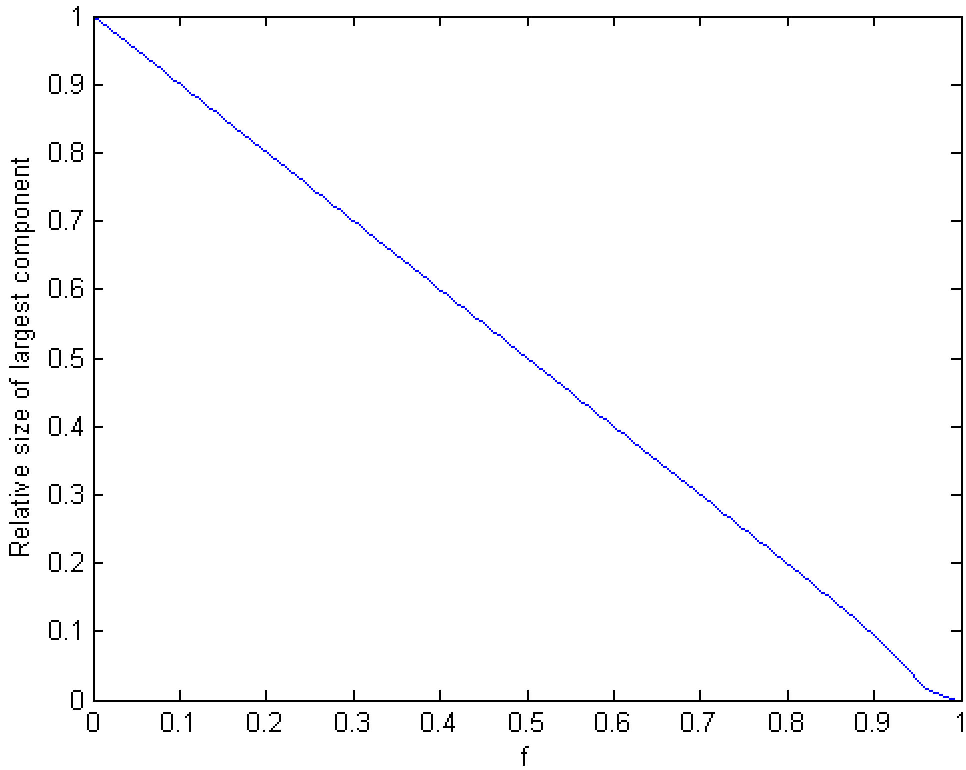

Fig. 5: Relative size of the largest cluster (S_1) as a function of the percentage (f) of nodes that are randomly removed from the Erdos-Renyi random network

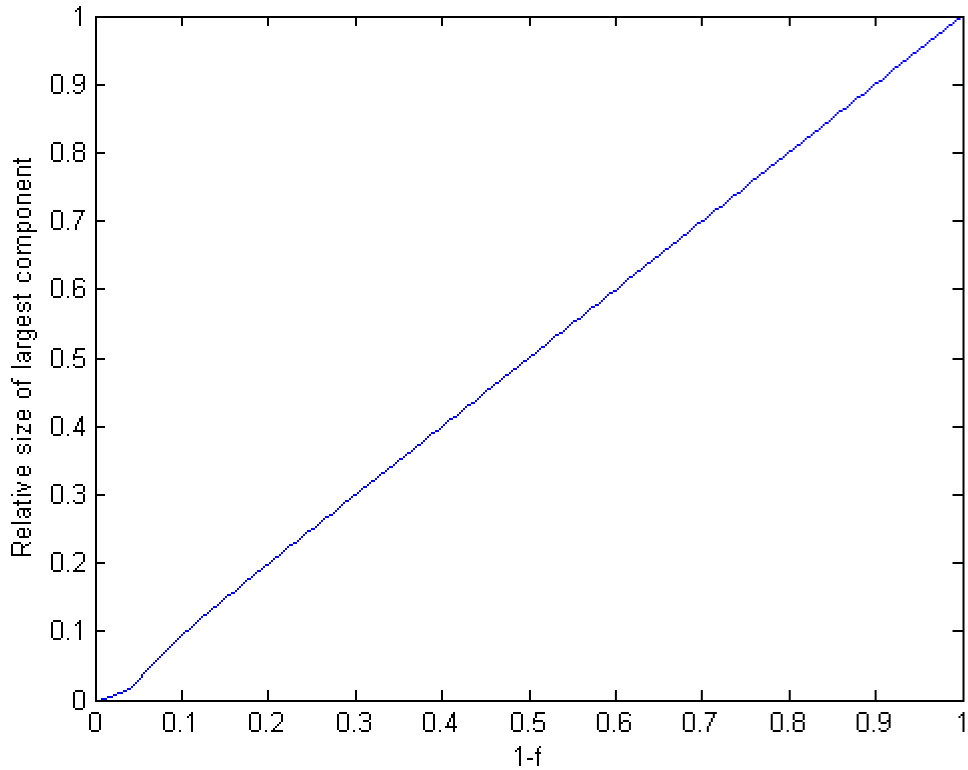

Fig. 6: Relative size of the largest cluster (S_1) as a function of the percentage (1-f) of nodes that are operational after random removal of (f) percent of nodes in the ER random network

4.2)

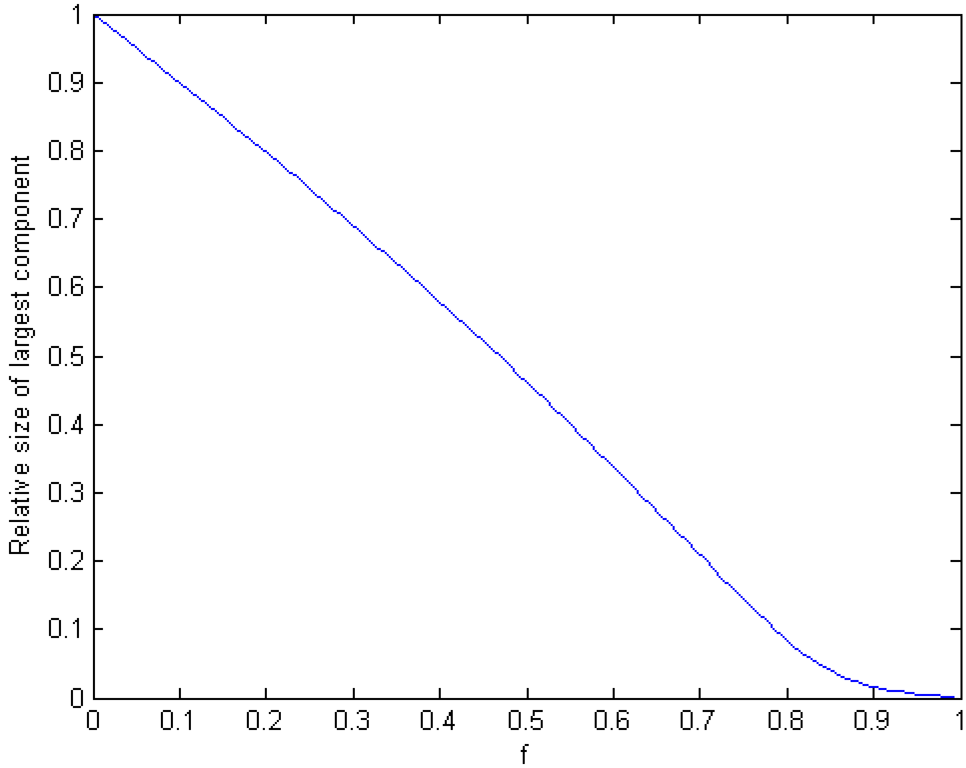

Fig. 7: Relative size of the largest cluster (S_1) as a function of the percentage (f) of nodes that are randomly removed from the BA scale free network

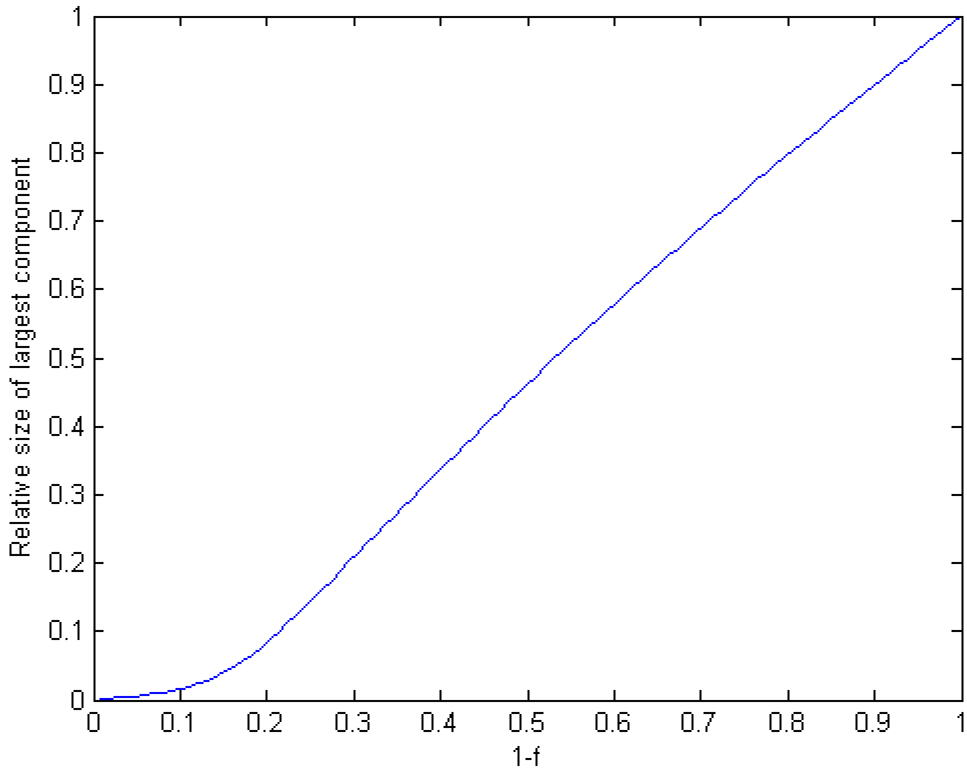

Fig. 8: Relative size of the largest cluster (S_1) as a function of the percentage (1-f) of nodes that are operational after random removal of (f) percent of nodes in the BA scale free network

4.3)

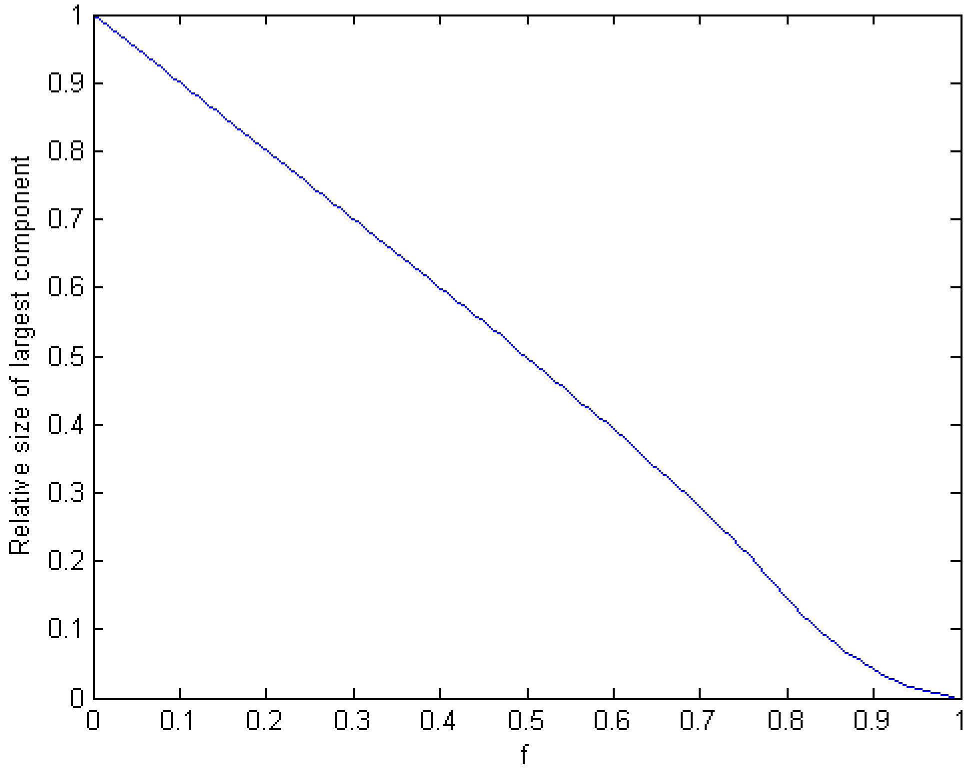

Fig. 9: Relative size of the largest cluster (S_1) as a function of the percentage (f) of nodes that are randomly removed from the Dutch Soccer network

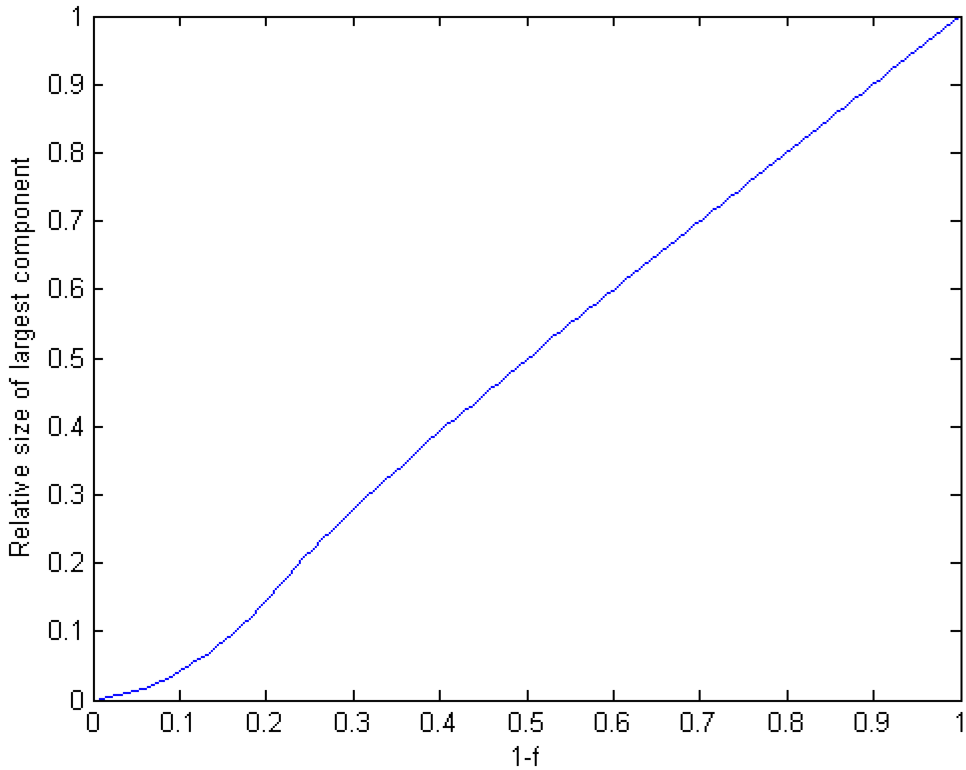

Fig. 10: Relative size of the largest cluster (S_1) as a function of the percentage (1-f) of nodes that are operational after random removal of (f) percent of nodes in the Dutch Soccer Network

4.4) **Robustness and percolation properties of the above network models:** The robustness of a network model is determined by the percolation transition which in turn is determined by the percolation threshold ($1-f_c$) and the growth of the largest component with increasing ($1-f_c$). Percolation threshold is a critical value of (1-f) at which a giant cluster or largest component appears in the corresponding percolation graph/plot. According to the figures 6, 8 and 10, Erdos-Renyi random graph has better percolation transition with lower percolation threshold and high growth rate of largest component with respect to (1-f) values. It is therefore more robust. Both, the Erdos-Renyi random network and the Dutch soccer network are more robust than the Barabasi-Albert scale free network. The relative robustness scale of the three network models is: **B-A scale free < Dutch Soccer < Erdos-Renyi**.

The reason for weak robustness of above scale free networks (i.e. both B-A and Dutch Soccer networks) compared to the above E-R random network is due to the presence of high degree nodes (hubs) in the scale free networks whose number is not very low compared to the number of low degree nodes in those networks. This phenomenon is due to the lower value of N in those network. So, in such networks during the random removal/failure of nodes in the network, the probability of removing hubs also becomes higher. As, hubs connect large parts of the network their removal can make the largest component disappear and will further disconnect the network too. Thus, above scale free networks are less robust compared to the above E-R random network. Further, higher robustness of Dutch Soccer network compared to the B-A scale free network can be explained due to the fact that both those networks have different degree distributions (i.e different power exponents) and preferential attachment. Furthermore, this behaviour of above network classes to percolation can also be accounted for pseudo random nature of the tool (i.e. Matlab) used.

For the Erdos-Renyi random network the average degree (E[D]) is 30. For $N \rightarrow \infty$, the theoretical percolation threshold ($1-f_c$) = $\frac{1}{E[D]}$ = 0.033. According to the figure 6, the percolation threshold value of the E-R random network as in section 4.1 is slightly less than the corresponding theoretical value ($1-f_c$) $\prec 0.033$. The reason being that the theoretical value calculated during the lecture is for large values of N (e.g. N = 100000) but the E-R random network as in section 4.1 has comparably smaller value of N = 685.

For the scale-free networks the threshold ($1-f_c$) = 0, if $N \rightarrow \infty$ and $2 \prec \gamma \prec 3$. According to the figure 8, the percolation threshold value of the B-A scale free network as in section 4.2 is larger than the corresponding theoretical value ($1-f_c$) $\succ 0$. Further, the curve also does not grow quickly with increasing $1-f_c$ values. The reason being that the theoretical value calculated during the lecture is for large values of N (e.g. N = 100000) and for $2 \prec \gamma \prec 3$ but the B-A scale free network as in section 4.1 has comparably smaller value of N = 685 and $\gamma \prec 2$.

*Note: N here is the number of nodes, f is the percentage/fraction of nodes removed from the network and γ is the power exponent of the scale free network.

4.5)

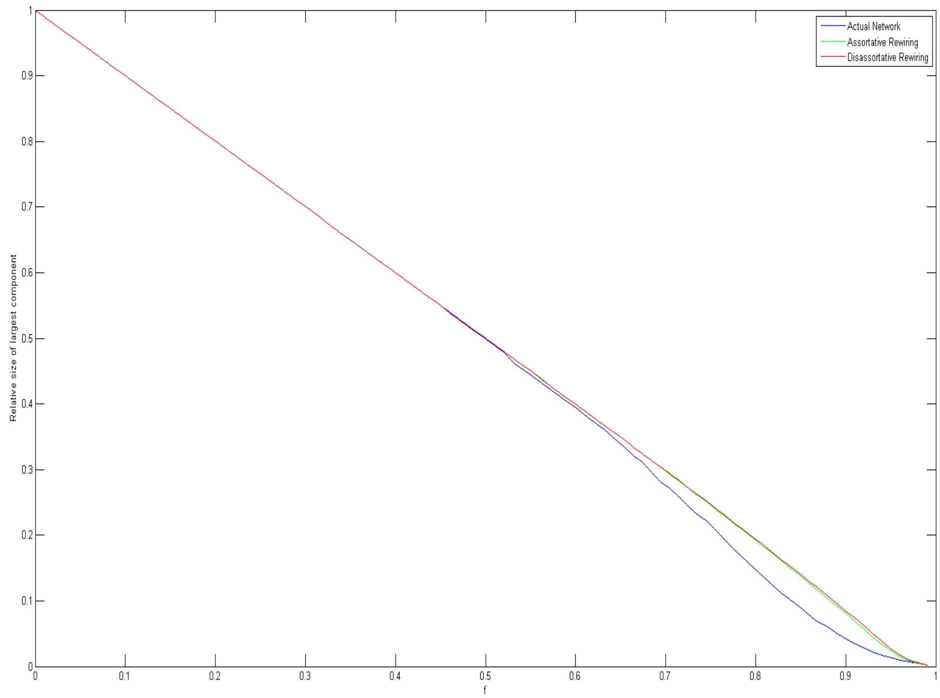

Fig. 11: Relative size of the largest cluster (S_1) as a function of the percentage (f) of nodes that are randomly removed from the three classes of networks as in sections 1, 2.1 and 2.2 (i.e. Dutch Soccer network in blue, assortative rewired Dutch Soccer network with $\rho \in [0.3, 0.31]$ in green and disassortative rewired Dutch Soccer network with $\rho \in [-0.31, -0.3]$ in red)

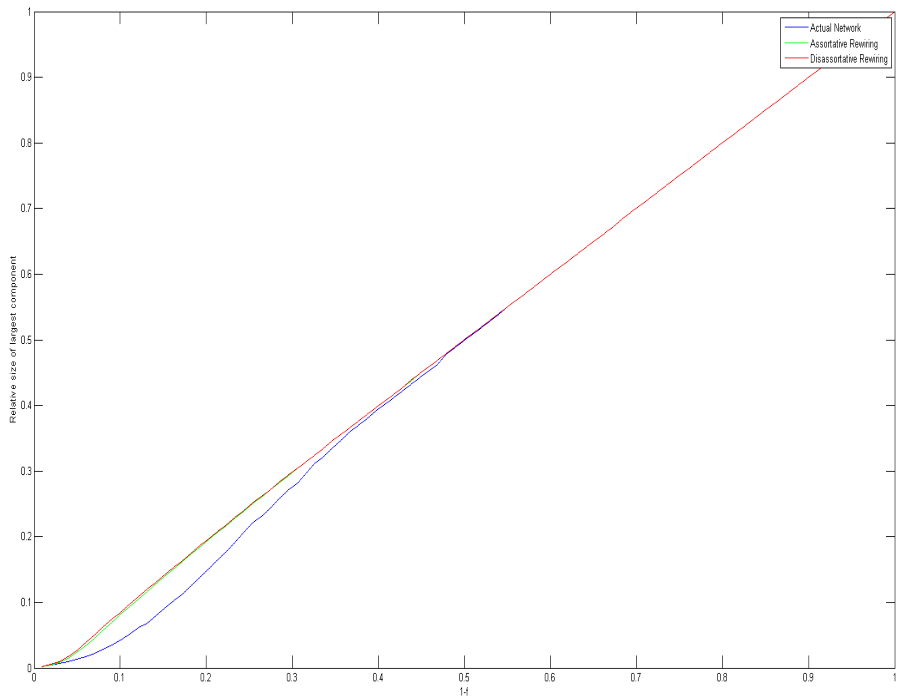

Fig. 12: Relative size of the largest cluster (S_1) as a function of the percentage (1-f) of nodes that are operational after random removal of (f) percent of nodes in the three classes of networks as in sections 1, 2.1 and 2.2 (i.e. Dutch Soccer network in blue, assortative rewired Dutch Soccer network in green and disassortative rewired Dutch Soccer network with in red)

The robustness of a network model is determined by the percolation transition which in turn is determined by the percolation threshold ($1-f_c$) and the growth of the largest component with increasing ($1-f_c$). According to the figure 12, rewired networks have better percolation transition compared to the actual network. The rewired networks have almost similar percolation transition but the disassortative rewired network has better percolation compared to the assortative rewired network. Thus, disassortative rewired network is the most robust network.The relative robustness scale of the three network models is: **Dutch Soccer < Assortative rewired < Disassortative rewired.**

The reason for disassortative rewired networks being more robust than the assortative rewired networks against random failures/removal can be explained by referring to the section 2.3. From the table in section 2.3, disassortative rewired network has slightly higher value of μ_{N-1} which translates to higher difficulty to disconnect the network into smaller disjoint clusters under random removal of nodes/links. Thus, disassortative rewired networks are slightly more robust compared to assortative rewired networks in the case of random node/link removal. The converse is true for targeted attacks/removal of nodes/links.

5 Opinion interactions

*Note: Below plots are obtained as an average over 100 iterations/instances

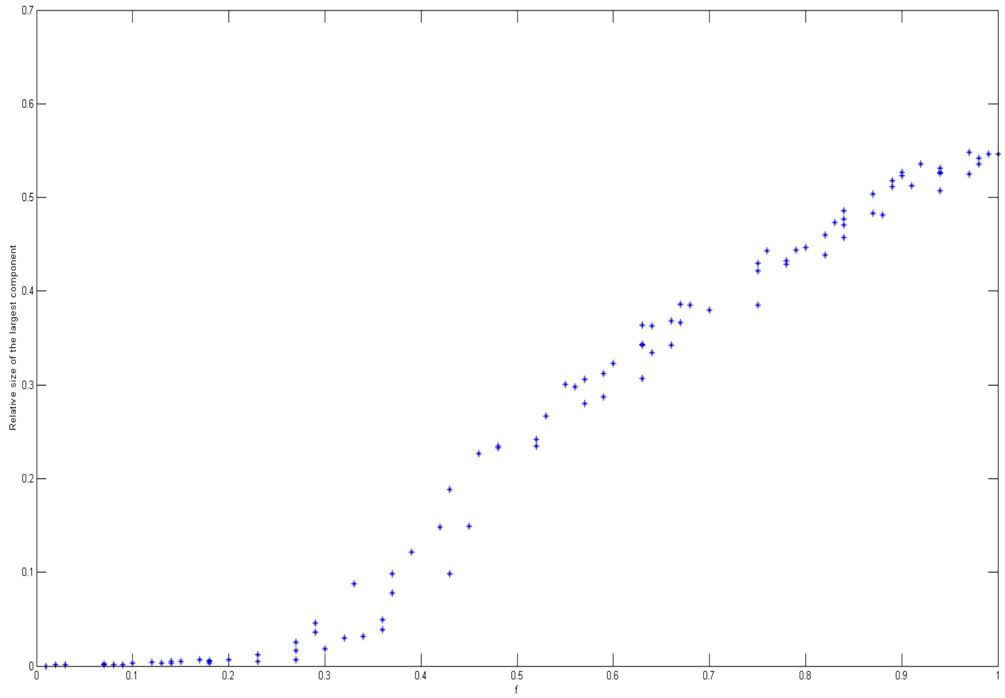

Fig. 13: Relative size of the largest cluster (S_1) of nodes with positive opinion in the study state as a function of the percentage (f) of nodes initially with positive opinion in an Erdos-Renyi random network with N = 2000 and *E*[D] = 4

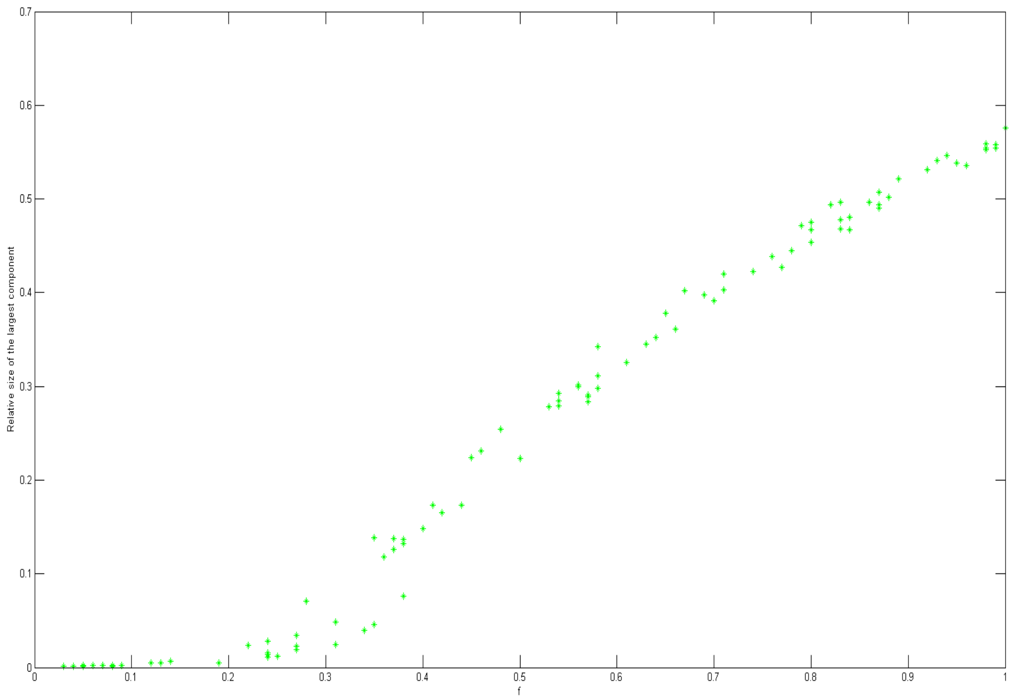

Fig. 14: Relative size of the largest cluster (S_1) of nodes with positive opinion in the study state as a function of the percentage (f) of nodes initially with positive opinion in an assortative rewired Erdos-Renyi random network with N = 2000 and E[D] = 4

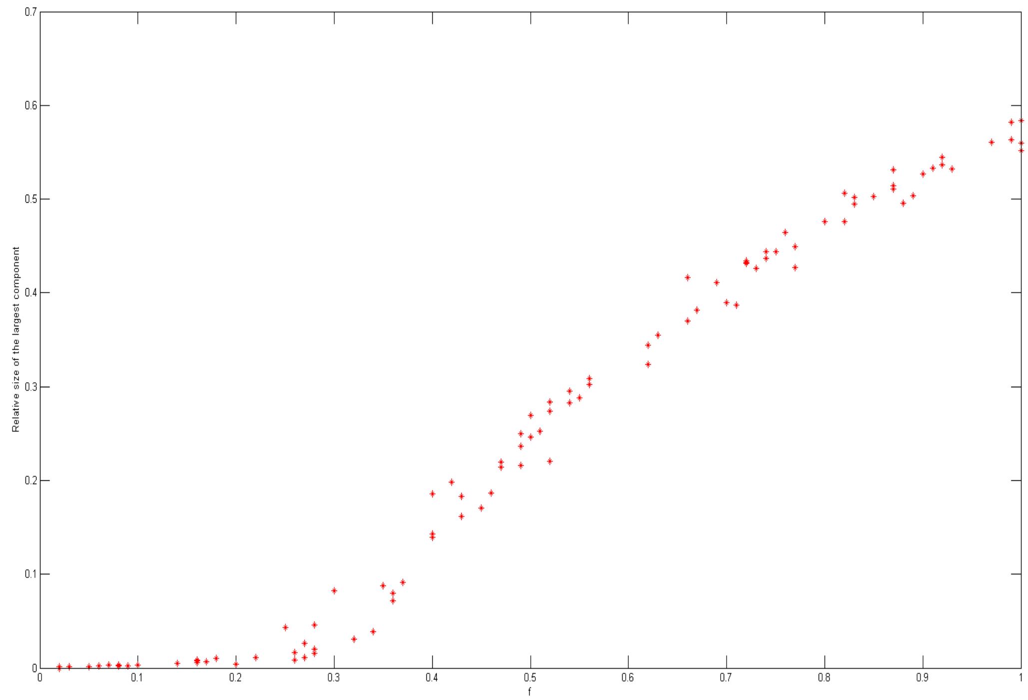

Fig. 15: Relative size of the largest cluster (S_1) of nodes with positive opinion in the study state as a function of the percentage (f) of nodes initially with positive opinion in an disassortative rewired Erdos-Renyi random network with N = 2000 and *E*[D] = 4

The minority opinion is more likely to survive in the case of normal Erdos-Renyi random network (fig. 13). This can be explained by observing the opinion threshold in the above graphs. Greater the threshold more chances of minority opinion survival in the network. Thus comparing threshold level in the above figures (i.e. fig. 13, 14 and 15) E-R random network has larger value of threshold compared to its rewired counterparts. Further, survival of minority opinion in an network can be determined by the value of λ_1 which is inverse of the dynamic processes (e.g. spread of epidemics, opinion interactions, etc.) threshold applied on the network.